Land Information System (LIS) Traceability Matrix

Submitted under Task Agreement GSFC-CT-2

Cooperative Agreement Notice (CAN)
CAN-00OES-01

Increasing Interoperability and Performance of Grand Challenge Applications in the Earth, Space, Life, and Microgravity Sciences

August 12, 2003

Revision 3.0

History:

Revision	Summary of Changes	Date
3.0	Milestone "I" Updates	August 12, 2003
2.0	Milestone "H" Updates	March 14, 2003
1.0	Initial Version (Milestone "E")	July 2002

1 Introduction

The purpose of this document is to provide a traceability matrix for the Land Information System (LIS) to be implemented under funding from NASA's Computational Technologies (formerly High Performance Computing and Communications) Project.

This traceability matrix is a table that maps Requirement numbers to Test Case numbers, and, in particular, provides the following information:

- Requirement Specification Number: The requirement paragraph number as listed in the Requirements Specification document.
- Requirement Statement: A brief paraphrase of the actual requirement as it appears in the Requirements Specification document.
- Subsystem: The particular subsystem of LIS that addresses the requirement. The subsystems of LIS are:
  - Main: consists of the main driver and land surface models.
  - Data Management
  - UI: User Interface
  - System Management: consists of tools for managing LIS's Linux cluster.
- Test Case #: The test case or procedure number of the test that will be run to verify the requirement.
- Verification: How well the requirement was verified: not verified; partially verified; fully verified
- Modification Field: Used in case requirement has been modified in any way throughout the life of the project.

2 Description

2.1 Land Surface Modeling and Data Assimilation

In general, land surface modeling seeks to predict the terrestrial water, energy and biogeochemical processes by solving the governing equations of the soil-vegetation-snowpack medium. Land surface data assimilation seeks to synthesize data and land surface models to improve our ability to predict and understand these processes. The ability to predict terrestrial water, energy and biogeochemical processes is critical for applications in weather and climate prediction, agricultural forecasting, water resources management, hazard mitigation and mobility assessment.

In order to predict water, energy and biogeochemical processes using (typically 1-D vertical) partial differential equations, land surface models require three types of inputs: 1) initial conditions, which describe the initial state of land surface; 2) boundary conditions, which describe both the upper (atmospheric) fluxes or states also known as "forcings" and the lower (soil) fluxes or states; and 3) parameters, which are a function of soil, vegetation, topography, etc., and are used to solve the governing equations.

2.2 Land Data Assimilation System (LDAS)

LDAS is a model control and input/output system (consisting of a number of subroutines, modules written in Fortran 90 source code) that drives multiple offline one dimensional land surface models (LSMs) using a vegetation defined "tile" or "patch" approach to simulate sub-grid scale variability. The one-dimensional LSMs such as CLM and Noah, which are subroutines of LDAS, apply the governing equations of the physical processes of the soil-vegetation-snowpack medium. These land surface models aim to characterize the transfer of mass, energy, and momentum between a vegetated surface and the atmosphere.

LDAS makes use of various satellite and ground based observation systems within a land data assimilation framework to produce optimal output fields of land surface states and fluxes. The LSM predictions are greatly improved through the use of a data assimilation environment such as the one provided by LDAS. In addition to being forced with real time output from numerical prediction models and satellite and radar precipitation measurements, LDAS derives model parameters from existing topography, vegetation and soil coverages. The model results are aggregated to various temporal and spatial scales, e.g., 3 hourly, $1/4°$.

The execution of LDAS starts with reading in the user specifications. The user selects the model domain and spatial resolution, the duration and timestep of the run, the land surface model, the type of forcing from a list of model and observation based data sources, the number of "tiles" per grid square (described below), the soil parameterization scheme, reading and writing of restart files, output specifications, and the functioning of several other enhancements including elevation correction and data assimilation.

The system then reads the vegetation information and assigns subgrid tiles on which to run the one-dimensional simulations. LDAS runs its 1-D land models on vegetation-based "tiles" to simulate variability below the scale of the model grid squares. A tile is not tied to a specific location within the grid square. Each tile represents the area covered by a given vegetation type.

Memory is dynamically allocated to the global variables, many of which exist within Fortran 90 modules. The model parameters are read and computed next. The time loop begins and forcing data is read, time/space interpolation is computed and modified as necessary. Forcing data is used to specify boundary conditions to the land surface model. The LSMs in LDAS are driven by atmospheric forcing data such as precipitation, radiation, wind speed, temperature, humidity, etc., from various sources. LDAS applies spatial interpolation to convert forcing data to the appropriate resolution required by the model. Since the forcing data is read in at certain regular intervals, LDAS also temporally interpolates time average or instantaneous data to that needed by the model at the current timestep. The selected model is run for a vector of "tiles", intermediate information is stored in modular arrays, and output and restart files are written at the specified output interval.

2.3 Community Land Model (CLM)

CLM is a 1-D land surface model, written in Fortran 90, developed by a grassroots collaboration of scientists who have an interest in making a general land model available for public use. LDAS currently uses CLM version 2.0, formerly known as Common Land Model. CLM version 2.0 was released in May 2002 and was implemented in LDAS in March 2003. The source code for CLM 2.0 is freely available from the National Center for Atmospheric Research (NCAR) (`http://www.cgd.ucar.edu/tss/clm/`). The CLM is used as the land model for the community climate system model (CCSM) (`http://www.ccsm.ucar.edu/`) and the community atmosphere model (CAM) (`http://www.cgd.ucar.edu/cms/`). CLM is executed with all forcing, parameters, dimensioning, output routines, and coupling performed by an external driver of the user's design (in this case done by LDAS). CLM requires pre-processed data such as the land surface type, soil and vegetation parameters, model initialization, and atmospheric boundary conditions as input. The model applies finite-difference spatial discretization methods and a fully implicit time-integration scheme to numerically integrate the governing equations. The model subroutines apply the governing equations of the physical processes of the soil-vegetation-snowpack medium, including the surface energy balance equation, Richards' [4] equation for soil hydraulics, the diffusion equation for soil heat transfer, the energy-mass balance equation for the snowpack, and the Collatz et al. [2] formulation for the conductance of canopy transpiration.

2.4 The Community Noah Land Surface Model

The community Noah Land Surface Model is a stand-alone, uncoupled, 1-D column model freely available at the National Centers for Environmental Prediction (NCEP; `ftp://ftp.ncep.noaa.gov/pub/gcp/ldas/noahlsm/`). Noah can be executed in either coupled or uncoupled mode. It has been coupled with the operational NCEP mesoscale Eta model [1] and its companion Eta Data Assimilation System (EDAS) [5], and the NCEP Global Forecast System (GFS) and its companion Global Data Assimilation System (GDAS). When Noah is executed in uncoupled mode, near-surface atmospheric forcing data (e.g., precipitation, radiation, wind speed, temperature, humidity) is required as input. Noah simulates soil moisture (both liquid and frozen), soil temperature, skin temperature, snowpack depth, snowpack water equivalent, canopy water content, and the energy flux and water flux terms of the surface energy balance and surface water balance. The model applies finite-difference spatial discretization methods and a Crank-Nicholson time-integration scheme to numerically integrate the governing equations of the physical processes of the soil vegetation-snowpack medium, including the surface energy balance equation, Richards' [4] equation for soil hydraulics, the diffusion equation for soil heat transfer, the energy-mass balance equation for the snowpack, and the Jarvis [3] equation for the conductance of canopy transpiration.

2.5 Variable Infiltration Capacity Model

Variable Infiltration Capacity (VIC) model is a macroscale hydrologic model, written in C, being developed at the University of Washington, and Princeton University. The VIC code repository along with the model description and source code documentation is available from `http://hydrology.princeton.edu/research/lis/index.html`. VIC is used in macroscopic land use models such as SEA - BASINS (`http://boto.ocean.washington.edu/seasia/intro.htm`). VIC is a semi-distributed, grid-based hydrological model, which parameterizes the dominant hydrometeorological processes taking place at the land surface - atmospheric interface. The execution of VIC model requires preprocessed data such as precipitation, temperature, meteorological forcing, soil and vegetation parameters, etc. as input. The model uses three soil layers and one vegetation layer with energy and moisture fluxes exchanged between the layers. The VIC model represents surface and subsurface hydrologic processes on a spatially distributed (grid cell) basis. Partitioning grid cell areas to different vegetation classes can approximate sub-grid scale variation in vegetation characteristics. VIC models the processes governing the flux and storage of water and heat in each cell-sized system of vegetation and soil structure. The water balance portion of VIC is based on three concepts:
1) Division of grid-cell into fraction sub-grid vegetation coverage.
2) The variable infiltration curve for rainfall/runoff partitioning at the land surface.
3) A baseflow/deep soil moisture curve for lateral baseflow.

Water balance calculations are preformed at three soil layers and within a vegetation canopy. An energy balance is calculated at the land surface. A full description of algorithms in VIC can be found in the references listed at the VIC website.

3 Traceability Matrix

Req. Spec. Number	Req. Statement	Subsystem	Test Case #	Verification	Mod. Field
3.1	GrADS-DODS for Data Management	Data Man.	-	-	-
3.2	LDAS for Data Assimilation	Main	-	-	-
3.3	CLM in LIS	Main	-	-	-
3.4	NOAH in LIS	Main	-	-	-
3.5	VIC in LIS	Main	-	-	-
3.6	ALMA for Input Variables	Main	-	-	-
3.7	ALMA for Output Variables	Main	-	-	-
3.8	ESMF Compliance	Main	-	-	-
3.9	Internet-enabled User Interface	UI	-	-	-
4.1	Land Surface Modeling	Main	-	-	-
4.2	Water and Energy Balance	Main	-	-	-
4.2.1	Computation at User-defined Time Intervals	Main	-	-	-
4.2.2	Mass and Energy Conservation	Main	-	-	-
4.3	Land/Water Mask	Main	-	-	-
4.4	Run-time Definition of Domain	Main	-	-	-
4.4.1	Domain Definition	Main	-	-	-
4.4.2	Dynamic Tile Use	Main	-	-	-
4.4.3	Tile Definition	Main	-	-	-
4.4.4	Time-stepping	Main	-	-	-
4.4.5	I/O of Gridded and Point Data	Main	-	-	-
4.4.6	Support for Time-dependent Variables	Main	-	-	-
4.4.7	Restart Support	Main	-	-	-
4.4.8	Start-time and End-time	Main	-	-	-
4.4.9	Mandatory Output	Main	4.3.4	Fully Verified	-
4.4.10	Output Frequency	Main	-	-	-
4.4.11	6-d Gridded Output	Main	-	-	Deleted
4.4.12	Quality Control Output	Main	-	-	-
4.4.13	Concurrently Running Ensembles	Main	-	-	New
4.4.14	Tile-space Output	Main	-	-	New
5.1	1 ms per grid cell per day Throughput	Main	4.2.4, 4.2.5	Fully Verified	-
5.2	0.4 ms per grid cell per day Throughput	Main	4.4.5, 4.4.6, 4.4.7	-	-
5.3	Performance Monitoring	Sys. Man.	-	-	-

Req. Spec. Number	Req. Statement	Subsystem	Test Case #	Verification	Mod. Field
6.1	User Levels	UI	-	-	-
6.2	Web Browser User Interface	UI	-	-	-
6.2.1	Read-only Access for General Public	UI	-	-	-
6.2.1.1	Animated or Still Output Images	UI	-	-	-
6.2.1.2	Contour or Shaded Output Images	UI	-	-	-
6.2.2	Password-restricted Access to Data	UI	-	-	-
6.2.2.1	Near-real-time Access to Data	UI	-	-	-
6.2.3	Password-restricted Access to Run Land Surface Models	UI	-	-	-
6.3	Configuration	UI	-	-	-
6.4	Initialization via Restart	Main	-	-	-
6.5	Write Restart Data	Main	-	-	-
6.6	Queuing System	Sys. Man.	-	-	-
6.7	Batch Mode for Operation	Sys. Man.	-	-	-
6.8	Debug Mode	Sys. Man.	-	-	-
6.9	Error Logging	Sys. Man.	-	-	-
6.10	Publicly Released Documentation and Source Code	UI	-	-	-
7.1	LIS Shall Run on LIS Cluster	Main	-	-	-
7.2	NOAH and CLM at 1/4 deg on SGI Origin 3000	Main	Baseline	Fully Verified	-
7.3	NOAH and CLM at 5 km on SGI Origin 3000	Main	4.2.1, 4.2.2, 4.2.3	Fully Verified	-
7.4	VIC on SGI Origin 3000	Main	4.3.1	Fully Verified	Modified
7.5	LDAS with Noah and CLM at 5 km on LIS Linux cluster	Main	4.3.2, 4.3.3	Partially Verified	Modified
7.6	LDAS and VIC on LIS Linux cluster	Main	4.3.1	Fully Verified	New
7.7	LDAS and LSMs at 1 km on LIS Linux cluster	Main	4.4.2, 4.4.3, 4.4.4	-	New Number
7.8	GUI Web Browser for User Interface	UI	-	-	New Number

Req. Spec. Number	Req. Statement	Subsystem	Test Case #	Verification	Mod. Field
8.1	Data Management Shall Support LIS	Data Man.	-	-	-
8.2	I/O in GrADS-DODS Format	Data Man.	-	-	-
8.3	Input Data	Main	-	-	Modified
8.3.1	Input Data Sources	Data Man.	-	-	-
8.3.2	Re-mapping of Input Data	Main	-	-	-
8.3.3	Re-projecting of Input Data	Data Man.	-	-	-
8.3.4	Input Data Spatial Interpolation	Main	-	-	-
8.3.5	Input Data Temporal Interpolation	Main	-	-	-
8.4	Output data	Data Man.	-	-	-
8.4.1	Output Data Formats	Data Man.	-	-	Modified
8.4.2	Output Data Conversion	Data Man.	-	-	-
8.4.3	Goode Homolosine for Output Data Projection	Main	-	-	Deleted
8.4.4	Lat/Lon for Output Data Projection	Data Man.	-	-	New
8.4.5	Re-projection of Output Data	Data Man.	-	-	New Number
8.5	Data Catalog	Data Man.	-	-	-
8.6	Automatic Update to Catalog	Data Man.	-	-	-
8.7	Backup of Data	Data Man.	-	-	-
8.8	Data Storage	Data Man.	-	-	-
9.1	Data Reliability	Data Man.	-	-	-
9.2	Authentication and Authorization Enforcement	Sys. Man.	-	-	-
9.3	Web Access Monitoring	Sys. Man.	-	-	-
9.4	Ftp Monitoring	Sys. Man.	-	-	-
9.5	Usage Limited	Sys. Man.	-	-	-
10.1	On-line Overview and Help	UI	-	-	-
10.2	FAQ	UI	-	-	-
10.3	Highlights Page	UI	-	-	-
10.4	On-line Tutorial	UI	-	-	-
10.5	User's Guide	UI	-	-	-

References

[1] F. Chen, K. Mitchell, J. Schaake, Y. Xue, H. Pan, V. Koren, Y. Duan, M. Ek, and A. Betts. Modeling of land-surface evaporation by four schemes and comparison with fife observations. *J. Geophys. Res.*, 101(D3):7251–7268, 1996.

[2] G. J. Collatz, C. Grivet, J. T. Ball, and J. A. Berry. Physiological and environmental regulation of stomatal conducatance: Photosynthesis and transpiration: A model that includes a laminar boundary layer. *Agric. For. Meteorol.*, 5:107–136, 1991.

[3] P. G. Jarvis. The interpretation of leaf water potential and stomatal conductance found in canopies of the field. *Phil. Trans. R. Soc.*, B(273):593–610, 1976.

[4] L. A. Richards. Capillary conduction of liquids in porous media. *Physics*, 1:318–333, 1931.

[5] E. Rogers, T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, M. Baldwin, N. W. Junker, and Y. Lin. Changes to the operational "early" eta analysis/forecast system at the national centers of environmental prediction. *Wea. Forecasting*, 11:391–413, 1996.